

Get the Crowd Going – An Empirical Analysis toward a Crowdsourcing Engagement Model

Nargess Tahmasbi
Penn State University
nvt5061@psu.edu

Henrik Sternberg
Iowa State University
henrik@istate.edu

Abstract

Crowdsourcing continues to attract attention from researchers, organizations, and policy makers alike. In particular, crowdsourcing projects that engage intrinsically motivated volunteers and are aimed at endeavors such as policy making, research, and social activism, need to understand how to create sustained engagement in their initiatives. A fair amount of research has been done on identifying users' motivation in crowdsourcing. However, crowdsourcing motivation literature still lacks sufficient theory-driven approaches. In this paper, we derive from theory of work motivation, the technology acceptance model, and the gamification concept to propose a model that can explain participants' motivation in crowdsourcing. To measure our model constructs, we conducted a survey on the users of a European crowdsourcing project, researching truck traffic tracking. The results of the exploratory factor analysis confirm that our constructs can be measured properly using our questionnaire. At the end of this paper, we explain our findings and the contribution of our study.

1. Introduction

Crowdsourcing refers to the use of technologies to gather the collective effort and wisdom from an undefined group of online users for organizational innovation and/or problem solving [1]. A large portion of research has addressed crowdsourcing in micro-tasking and creative work [2]. Some research has investigated community crowdsourcing, a community in which participants collaborate on larger projects, such as open source software development [3]. Given the successful outcome of crowdsourcing in various initiatives, researchers and policy makers alike are paying attention to the role of crowdsourcing policy work and/or work that aims to support policy making [4] and the adjacent fields of social and digital activism [5].

Motivation and engagement are tightly coupled, as intrinsic and extrinsic motivation determines the level of engagement in the community [6]. Many crowdsourcing

studies strive to increase the participants' engagement [7], [8]. A recent review by Mindel et al. suggests polycentric principles to aid participation in community crowdsourcing [9]. However, research is sparse on different types of crowds and their respective motivation and engagement. This is even more so for crowdsourcing aiming at policy change, as most such crowds fail to gain traction, legitimacy and sustained engagement [4].

Further understanding the engagement of crowds is of great value not only to researchers in information systems (IS), but also to public policy makers, leaders of social movements (e.g., representatives of unions and associations) and communities, and to researchers in various fields where crowdsourcing can open up to novel datasets. Hence, this paper sets out to answer the research question:

What factors contribute to sustained engagement in crowdsourced research for societal change?

The aim is to contribute to theory on antecedents for crowdsourcing user engagement, in the context of research and policy-making. Studies have suggested some of the potential motivational factors that act as antecedents for engagement in non-competitive, pro-bono crowdsourcing [12]. We seek to adapt from the previous models and provide insights that can explain the participation engagement in this particular subset of crowdsourcing.

Participation engagement is the main phenomenon of interest in this study. It is mentioned in the literature from different perspectives, but the research has not been able to come up with a unified way to measure it. Meanwhile, studying motivational factors that lead to engagement requires a reliable definition of the term and metrics for measuring engagement in crowdsourcing. In this study, we use Nguyen et al.'s proposed definition and their quantitative measures of engagement [13]. They define crowdsourcing engagement as a metric that comprises four dimensions: “*magnitude, temporal intensity, diversity, and recency of tangible effort*” [13]. The context of their study was a community crowdsourcing platform for citizen engagement that shares similarities with the context of this study. Engagement is the only quantitative data-driven

construct that will be measured using the users' activity. However, in this study, our focus is to ensure that the qualitative constructs of the model are measurable.

We use the crowdsourcing project "Cabotagestudien" (Swedish for "The Cabotage study") to study engagement, in particular, from the data collected between 2013 and 2016. The Cabotage study was about mapping the movements of all trucks in Scandinavia and Austria by engaging some 8,000 volunteers who used a smartphone app and participated in the Cabotage study Facebook page. The engagement in, and the results of the project, sparked considerable public attention and debate, resulting in impacts on policy (e.g., [14]) and research on social issues in supply chain management (e.g., [15]).

This paper is structured as follows. In section 2, we provide background on the motivation studies in crowdsourcing and discuss the theoretical foundation related to the study. In section 3, we use previous theoretical foundations to hypothesize and present a model of crowdsourcing motivation in the context of crowdsourcing in social research. In section 4, we explain the dataset and the methodology followed by construct measurement methods. In section 5, we present and discuss the results of the exploratory factor analysis, the reliability of the model constructs, and the study's limitations. In section 6, we discuss the theoretical implications and we conclude the paper in section 7 by offering suggestions for future research.

2. Motivation and engagement

2.1. Crowdsourcing motivation

To increase engagement in crowdsourcing, studies seek to understand what motivates participants to engage in crowdsourcing tasks. Crowdsourcing motivation studies prevail in the crowdsourcing research. By adapting traditional motivation theories, many studies categorize such engagement into two classes of motivations: intrinsic and extrinsic [6], [16], [17]. Depending on the nature of crowdsourcing projects, one of these types of motivations may have a stronger presence over the other. For example, in a crowdsourcing project where participants are paid, extrinsic motivations, such as monetary incentives, are stronger than intrinsic ones. In the presence of extrinsic motivations, intrinsic ones, such as altruism, have a weaker effect on the task effort [6], [18].

2.2. Work motivation theory

Motivation is a phenomenon that has been studied in various contexts, and the study of work motivation has been around for decades. Hackman and Oldham introduced a work motivation theory that specifies three

interacting conditions that impact one's motivation to perform better on the job: the psychological states of employees, the characteristics of jobs, and the attributes of individuals that define their response to a challenging task [19]. The critical psychological states resulting from the job characteristics are: the perceived meaningfulness of the work, the experienced responsibility of the outcome, and the knowledge of the actual results of the work. These psychological states lead to personal and work outcomes. Crowdsourcing tasks by nature are different from tasks on the job, but the work motivation in the two is similar. A few crowdsourcing studies have adapted the work motivation theory to crowdsourcing [20].

2.3. Technology acceptance model

Davis introduced the technology acceptance model (TAM), suggesting that the perceived ease of use (PEU) and perceived usefulness (PUE) determine the intention to use a technology (particularly information systems) [21]. One of the fundamental components of crowdsourcing is technology [22]. Hence, participation in crowdsourcing tasks can be partially associated with the individuals' perceived ease of use and perceived usefulness of the technology. This technology can be the crowdsourcing platform or a mobile app that is being used as the main participation platform. Despite the important role of technology as an enabler of crowdsourcing, previous research has not investigated the implications of TAM on the success of crowdsourcing projects. In most crowdsourcing cases, especially the ones using a mobile platform, the technology is so intertwined with the task that it is infeasible to separate them from each other. In this respect, the perceived usefulness of the technology in TAM is more of a question pertaining to both technology and the task at the same time, to the extent that the usefulness of the technology can be indexed by measuring the meaningfulness of the crowdsourcing task.

2.4. Gamification in crowdsourcing

Gamification is a relatively new phenomenon compared to crowdsourcing. It is simply defined as "*the use of game design elements in non-game contexts*" [23]. Gamification intends to increase participants' motivation in non-gaming contexts and includes features such as points, leaderboards, badges, and bonus cards. A number of studies suggest that gamification has a positive effect on increased engagement in online programs that exclude crowdsourcing [24]. A few studies have also investigated the use and role of gamification in the success of different applications of crowdsourcing including tourism [25], the medical domain [26] and data collection [27]. While

gamification in general aims to provide a positive effect on engagement, its effect depends on the context in which the gamification is implemented [28]. Crowdsourcing projects employ gamification in different ways and depending on the approach they take, they get different gamification outcomes. For example, crowdsourcing projects with simple and monotonous tasks may employ simpler and less diverse gaming elements such as points or levels than crowdsourcing projects with a more diverse set of tasks [29].

3. Theory and hypotheses development

In the wake of a call for a theory of crowdsourcing engagement, we are exploring multitudes of factors that may affect the sustained user engagement in crowdsourcing. Hence in this section, our hypotheses are developed in a way that they introduce factors from several relevant theoretical foundations and place them as moderators in the model. Eventually, the model needs to go through cycles of testing and evaluation until the relative roles of the factors in the model are specified. Departing from the theoretical foundations discussed in the previous section, we hypothesize that five factors have significant impacts on users' engagement. de Vreede et al. discuss the effects of personal interest in the topic and goal clarity on motivation/engagement [12]. Moreover, work motivation theory suggests that the perceived meaningfulness of the work effects an individual's decision to perform well on the job [19]. An adaption of the previous findings implies that if the user is interested in the topic and has a positive attitude toward crowdsourcing work, he/she may be willing to try the technology and participate in the crowdsourced task. Hence, the first hypothesis:

H1: Personal interest and attitude toward crowdsourcing work is positively related to engagement in crowdsourcing.

Drawing on work motivation theory [19], we imply that the user's engagement is also affected by the amount and quality of the feedback that he/she receives and the extent to which he/she feels responsible for the outcome of the work. However, this is more of a moderating effect, because without personal interest and a positive attitude toward the topic, the user will not be motivated to participate. Whether or not he/she is willing to continue participation depends on the factors that are suggested by the work motivation theory [19]. Thus, the following hypothesis:

H2: Received feedback and perceived autonomy positively moderate the effects of personal interest and attitude on engagement.

The technology acceptance model suggests that the user's acceptance of a new information system depends

on two main factors: the perceived ease of use (PEU) and the perceived usefulness (PUE) [30]. We propose that the technology acceptance model (TAM) is relevant to crowdsourcing projects because the participants use one or another type of information system (technology). In addition to the social factors pertaining to crowdsourcing engagement, technological factors are also important in motivating a user to engage and contribute. Thus, even if the user is initially engaged and interested in crowdsourcing, if the technology is not easy to use (PEU), the user may not be willing to continue participating. Technology is the central phenomenon of interest in TAM. In our crowdsourcing study, though, the technology is peripheral to the task, not the main phenomenon. The task, of course, is introduced through the technology, which is a foundation of crowdsourcing [22], but is not central to it. At the same time, as mentioned before, the two are inevitably intertwined with each other and the separation of the two is not practical. Thus, as long as the crowdsourcing task is useful and meaningful, there is no doubt about the usefulness of technology for the task, but the design of the technology does impact its ease of use. Consequently, we did not include the usefulness of technology from TAM as a factor in our model, however, this factor is taken into consideration in the construct for personal interest and attitude toward the task that includes the perceived meaningfulness of the task. Moreover, we argue that perceived clarity of the goal of the task is also a relevant component in the user's perceived ease of use. When the goals are clear to the user, different components of the system are also perceived as meaningful in accordance with the goal they serve. Hence, the following hypothesis:

H3: Perceived goal clarity and ease of use of the crowdsourcing app positively moderates the effects of personal interest and attitude on engagement.

But since the task and the technology are such interconnected components in crowdsourcing [22], the task difficulty is also relevant to the ease of use (PEU) of the technology. Even if the user is highly interested in the topic, if the task demands lots of effort and commitment beyond the user's available resources, the user may not be motivated enough to continue participating. Hence, we propose that the perceived required task effort is another moderator of the causal relationship between personal interest and engagement. Accordingly, the following hypothesis is:

H4: Perceived required task effort negatively moderates the effect of personal interest and attitude on engagement.

As discussed in the literature, gamification can potentially increase the engagement rate in

crowdsourcing projects [28]. Gamification provides enjoyment, and the enjoyment experienced from a process can be one of the reasons people use a system [21]. Depending on the context, and the gamifying process, the experienced joy that resulted from the existence of gamification elements can play a role in motivating an interested user to engage in the crowdsourcing project via the technology. Hence, the final hypothesis is:

H5: Experienced joy from gamification elements positively moderates the effect of personal interest and attitude on engagement.

Figure 1 depicts a crowdsourcing engagement model for the Cabotage study based on the five hypotheses and the causal and moderating relationships among factors.

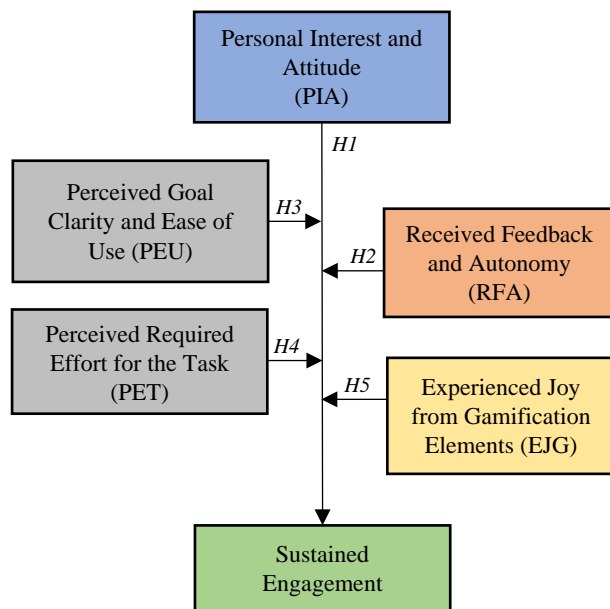


Figure 1. The crowdsourcing engagement model

4. Data and Methodology

We extracted primary data on the user activities in a crowdsourcing project (the Cabotage study) from the system database. To analyze the different factors that determine crowdsourcing engagement (our unit of analysis), we developed and adapted scales from the IS literature. These scales were tested in a pilot survey with randomly chosen users. In addition to surveying motivational factors, we also collected descriptive data on the characteristics of the participants, for example, union membership (Figure 3).

4.1. The cabotage study crowd

The volunteers engaged in the cabotage study mainly contributed by making observations using a smartphone app (also named “Cabotagestudien”) [15].

The purpose was to create a snapshot of all national and international truck traffic. When the users saw a truck, they used the app to report its license plate number, referred to as “tagging” the truck (See Figure 2 for screenshots of the app). All the aggregated observations of all trucks were published on the study website (www.cabotagestudien.se), but with their respective license plates anonymized. The website included a leaderboard; however, most of the discussion about the project was carried out in Facebook groups (one group for each language of the users: Swedish, Danish, Norwegian and German). The Facebook groups gave the users opportunities to ventilate their opinions about the project and the app.



Figure 2. Screenshots of the Cabotage Study mobile app

The number of observations (using the app) signifies the individual level of sustained engagement. For the purpose of this paper, we selected all accounts (two account types existed: Facebook login and email + password login) with more than 5 observations (tasks carried out), in order to exclude users who set up an account just to test the app. The entire population consisted of 6,250 users with between 5 and 19,831 observations. A random sample of 350 users was selected and sent an engagement survey (Appendix, under the column heading “Item description”). The people were contacted through either email or a Facebook message, depending on their login credentials. The invitation letter contained brief information about the purpose of the survey and information about a raffle and the chance to win a new smartphone. After two reminders and an announcement on the Facebook page of the study, 61 users had responded (16% response rate). Two further users were contacted by one of the authors and persuaded to answer the survey. The first non-respondent stated he just had not had the time and the second respondent did not answer because he felt that he had not contributed enough to voice his opinion. The two non-respondent’s responses did not differ from their responding

counterparts. Out of a total of 63 responses, we were not able to match 6 of the responses to a particular user (during the process of this work, Facebook account links became inaccessible). In other words, 57 usable responses remained.

Figure 3 depicts the occupation of the participants in the study. 67% were drivers and 24% belonged to the category “other”. In the survey, the respondents had a text field in which they could specify “other”, and all but one choose to do so. This revealed that they were, for example, students, mechanics, a chef, retired and a traffic controller.

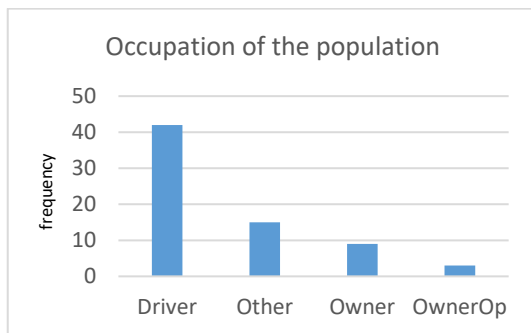


Figure 3. Occupation of the survey participants

4.2. Construct measurement

We adopted and adapted the measurement methods suggested by previous studies to measure most of our variables. The Appendix presents a complete listing of the constructs, items, indicators, and scales including their sources.

Personal interest and attitude (PIA) was measured using questionnaire items adapted from related questionnaire items in [31] and [32].

Received feedback (knowledge of the actual results) and perceived autonomy (sense of responsibility for the outcome) (RFA) were operationalized using the questions adapted from the study of flow in online shopping in [33] and from [34].

We used Davis’s TAM questionnaire to measure the perceived ease of use [30] and combined it with the perceived goal clarity measured by [33] in a single construct called perceived clarity and ease of use (PEU).

Task difficulty or the perceived required effort for the task (PET) was measured based on items from a study of multitasking and task difficulty [35].

Gamification elements were adapted from [29] and the questionnaire items were compiled that asked about the user’s experience of joy from those elements. This construct is called experienced joy from gamification elements (EJG).

We used Nguyen et al.’s approach to measure crowdsourcing engagement [13]. The main indicator of

an activity in the Cabotage study is tagging the trucks. There is not much diversity in the types of tasks one can perform in the app. The dataset includes the recorded tags and timestamps associated with each tagging activity for each user. Thus, it is feasible to measure the sustained engagement as suggested by Nguyen et al. [13]. This is the only construct of the model that we measured quantitatively, and it served as our dependent variable.

All other constructs were operationalized using questionnaire methods. Some of the measures from previous studies were on a 5-point Likert scale and others were on 7-point scale. To ensure the reliability of the original measures, we did not change the scales as originally developed by the previous studies while conducting the survey. However, after conducting the survey and collecting the answers, we converted the answers that were originally on 5-point Likert-scale to the 7-point Likert-scale to create a homogenous scale for Factor analysis and consequent analyses.

5. Results

We performed an exploratory factor analysis on the questionnaire results (57 observations) using factor analysis in R.

5.1. Exploratory Factor Analysis (EFA)

The scree graph suggests that about five factors (constructs) are sufficient for the factor analysis. We performed the exploratory factor analysis using the *factanal* function in R. The chi square statistic was 127.03 at 115 degrees of freedom. The p-value for the hypothesis that 5 factors are sufficient is 0.209, so we did not reject the hypothesis. The root mean square of the residuals (RMSR) was 0.05, and the df corrected root mean square of the residuals was 0.07. The fit, based upon off-diagonal values, was 0.97 which indicates a good fit (>0.95).

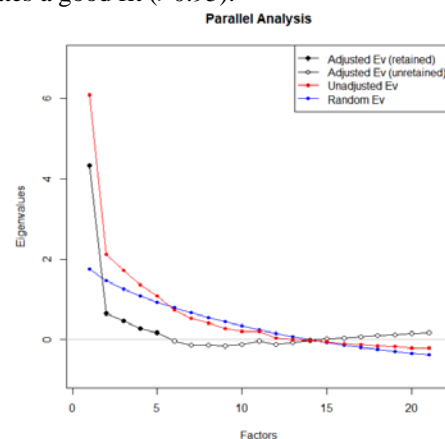


Figure 4. The Result of Parallel Analysis

To further validate the suitability of our 5-factor solution, we performed a parallel analysis on the data.

Figure 4 and Figure 5 present the results of the Horn's parallel analysis based on 5000 iteration using the mean estimate. The results from this analysis also suggests 5 factors to be retained.

Factor	Adjusted Eigenvalue	Unadjusted Eigenvalue	Estimated Bias
1	4.324691	6.074928	1.750236
2	0.648327	2.112136	1.463809
3	0.464300	1.719078	1.254778
4	0.274375	1.351380	1.077005
5	0.156914	1.081224	0.924309
6	-0.049606	0.734784	0.784390
7	-0.139070	0.519314	0.658384
8	-0.144100	0.398815	0.542916
9	-0.170421	0.263085	0.433506
10	-0.130468	0.201942	0.332410
11	-0.045999	0.192202	0.238201
12	-0.122290	0.027677	0.149968
13	-0.075347	-0.00716	0.068185
14	-0.022079	-0.03024	-0.00816
15	0.014334	-0.06377	-0.07810
16	0.031788	-0.11161	-0.14340
17	0.066060	-0.13664	-0.20270
18	0.097770	-0.15856	-0.25633
19	0.117347	-0.18794	-0.30529
20	0.142305	-0.20711	-0.34942
21	0.168607	-0.22342	-0.39203

Adjusted eigenvalues > 0 indicate dimensions to retain.
(5 factors retained)

Figure 5. Results of Horn's Parallel Analysis for factor retention

Table 1 shows the factor loadings that were achieved with a varimax rotation and a cutoff point of 0.4.

Table 1. The results of factor loadings

Item	Factor 1 (PEU)	Factor 2 (RFA)	Factor 3 (EJG)	Factor 4 (PIA)	Factor 5 (PET)
GoalDefined	0.68				
ClearInteract	0.66				
Skillful	0.66				
EaseUse	0.70				
DoingWell		0.71			
Perform		0.88			
ExpressIdea		0.60			
Control		0.50			
NewBadge			0.68		
TopLeader			0.98		
Points			0.58		
Reward			0.66		
Meaningful				0.85	
PosDiff				0.68	
WellBeing				0.43	
Important				0.41	
Complex	0.42				0.62
Demanding					0.55
Thoughts					0.68
TimeConsum					0.62
Challenging					0.49

The loading results show that the items appropriately load into the five factors with minimal cross-loadings. Factor 1 is related to goal clarity and ease of use. We call it perceived goal clarity and ease of use (PEU). Factor 2 combines feedback and perceived autonomy; we call it received feedback and perceived autonomy (RFA). The items in Factor 3 are all related to what we call the experienced joy from the gamification (EJG) elements of the task. Factor 4 includes the items that indicate the users' personal interest and attitude toward the task (PIA). Factor 5 measures the required effort for the task, as the list of items in this factor suggests. It is called perceived required effort for the task (PET). However, we observed one questionnaire item that cross-loaded into both factors 1 and 5. The loading of the *Complex* item for factor 1 was 0.42, and for factor 5 was 0.62. Since the loading of this item was strong enough in the correct factor (>0.6) and was significantly larger than its loading in the weak factor (with a difference ≥ 0.2), we were more inclined toward keeping this item loaded into factor 5. To further investigate improvement of this issue, we performed the factor analysis again with oblimin rotation this time. It produced similar factor loadings except that the *Complex* item did not have cross-loading and appropriately loaded into factor 5 (with a cut-off point of 0.4).

The correlations of the factors for the resulting 5 factors from oblimin rotation were all below 0.4, which indicates that the extracted factors are not significantly correlated. Table 2 presents the factor correlations.

Table 2. Factor correlations. (RFA: received feedback and perceived autonomy; EJG: experienced joy from gamification; PEU: perceived clarity and ease of use; PET: perceived required effort of the task; PIA: personal interest and attitude)

	RFA	EJG	PEU	PET	PIA
RFA	1.00				
EJG	0.35	1.00			
PEU	0.37	0.23	1.00		
PET	0.02	-0.03	0.06	1.00	
PIA	0.18	0.16	0.30	0.09	1.00

The proportion of variance explained by each factor is 24% for RFA, 21% for EJG, 22% for PEU, 16% for PET, and 17% for PIA. This makes RFA the strongest factor of them all.

5.2. Internal consistency reliability

After exploring the factors and the item loadings, we investigated the internal consistency reliability of the factors using Cronbach's alpha [36]. This test was used to confirm how well the items in each construct (factor) produced similar scores [36]. Table 3 presents the

Cronbach's alpha values for the five factors resulted from the EFA.

Table 3. Cronbach's alpha and internal consistency of the factors

Factor	Cronbach's alpha	Internal consistency
PIA	0.722	Acceptable
RFA	0.843	Good
PEU	0.827	Good
PET	0.704	Acceptable
EJG	0.842	Good

The literature suggests that a Cronbach's alpha between 0.8 and 0.9 is considered good and an alpha between 0.7 and 0.8 is acceptable [36]. The results suggest acceptable and good internal consistency for all factors.

6. Discussion

The exploratory factor analysis resulted in 5 factors with at least four items in each factor with empirically acceptable loadings. The internal consistency reliability of the items in each factor examined with Cronbach's alpha also suggests that the items properly measured each factor. Moreover, the factors do not show high correlations with each other, which indicates that they are significantly independent from each other. The above results suggest that the measurement method is reliable. Hence, the initial model is reliable and can be used as the final model for the next phase of our study. In the second pilot study, we will need to perform a confirmatory factor analysis to further test the reliability of the questionnaire with a larger sample size.

As highlighted in the introduction, this research is geared towards the context of the Cabotage study: crowdsourcing applied to research and/or social activism.

A major limitation of this research is the size of the pilot test: only 57 usable responses from a sample of 350 randomly selected users. Although the responses seemed to be quite homogenous (including the 6 the research team was unable to connect to a specific user's activity level), regardless of whether the volunteer user was a truck driver, owner-operator, fleet-owner or had a miscellaneous relation to the industry, it cannot be ruled out that a larger sample and a higher response rate is likely to generate results of higher significance. More specifically, the literature suggests that an exploratory factor analysis generates more generalizable results with larger sample sizes [37]. Thus, the next phase of the research will be a second pilot study in which we target a larger number of participants to increase the validity of the survey results.

In the sample, 21 of the 42 drivers were union members. The union drivers carried out on average 112

observations per user, whereas the non-unionized drivers carried out only 42 observations on average per user over the whole period. In the context of this research, this indicates that union members were more prone to engage in their industry.

The homogeneity of the answers from the sample is likely explained by the self-selection of the crowd participants, this being a volunteer activity.

7. Conclusion

This study set out to address the question of motivational factors in crowdsourcing engagement. We drew on multiple theoretical foundations in work motivation [19], a technology acceptance model [30], and gamification [29] to identify the influential factors that impact participants' engagement in crowdsourcing projects. We conducted a pilot study to investigate and confirm our measurement approach for the constructs that were proposed in our model. The results of the exploratory factor analysis were promising as the variables perfectly loaded into the factors as proposed. Moreover, the factor correlation results suggest that the extracted factors are not significantly correlated.

The analysis shows that RFA (Received Feedback and perceived Autonomy) and PEU (Perceived goal clarity and Ease of Use) are the strongest factors contributing to the sustained engagement of the crowd in this case. This is not surprising, as it confirms the importance of straight communication and clear feedback, which has been shown in previous literature on management and coordination of truck drivers [38], [39]. These conclusions are however weak, given the limited number of respondents and should be interpreted with that limitation in mind.

In the next phase, we will carry out a second pilot study to confirm the reliability of the questionnaire by testing it on a larger group of participants. Then we will proceed to the main study in which we run the refined survey and test the final version of the model using a regression analysis.

8. Acknowledgements

The authors are grateful to all the volunteers that participated in the study as well as to Isak Nyström, Lund University, for helping out with the data collection. Henrik would also like to express his gratitude to the Swedish Transport Union and the Lindley Foundation for financial support.

9. References

- [1] J. Howe, "Crowdsourcing: A definition," 2006. [Online]. Available: http://crowdsourcing.typepad.com/cs/2006/06/crowdsourcing_a.html. [Accessed: 06-Jun-2018].

- [2] Y. Huang, "Crowdsourcing New Product Ideas under Consumer Learning," *Management Science*, vol. 60, no. 9, pp. 2138–2159, 2014.
- [3] G. Von Krogh, S. Haefliger, S. Spaeth, and M. W. Wallin, "Carrots and rainbows: Motivation and social practice in open source software development," *MIS Quarterly*, pp. 649–676, 2012.
- [4] T. Aitamurto, H. Landemore, and J. Saldivar Galli, "Unmasking the crowd: participants' motivation factors, expectations, and profile in a crowdsourced law reform," *Information, Communication & Society*, vol. 20, no. 8, pp. 1239–1260, Aug. 2017.
- [5] L. Selander and S. L. Jarvenpaa, "Digital action repertoires and transforming a social movement organization," *MIS Quarterly*, vol. 40, no. 2, pp. 331–352, Jun. 2016.
- [6] J. Rogstadius, V. Kostakos, A. Kittur, B. Smus, J. Laredo, and M. Vukovic, "An Assessment of Intrinsic and Extrinsic Motivation on Task Performance in Crowdsourcing Markets," *ICWSM*, vol. 11, pp. 17–21, 2011.
- [7] M. Garcia Martinez, "Solver engagement in knowledge sharing in crowdsourcing communities: Exploring the link to creativity," *Research Policy*, vol. 44, no. 8, pp. 1419–1430, Oct. 2015.
- [8] A. Mao, E. Kamar, and E. Horvitz, "Why stop now? predicting worker engagement in online crowdsourcing," in *First AAAI Conference on Human Computation and Crowdsourcing*, 2013.
- [9] V. Mindel, L. Mathiassen, and A. Rai, "The Sustainability of Polycentric Information Commons," *Management Information Systems Quarterly*, vol. 42, no. 2, pp. 607–631, 2018.
- [10] E. Cullina, K. Conboy, and L. Morgan, "Measuring the crowd: a preliminary taxonomy of crowdsourcing metrics." In *Proceedings of the 11th International Symposium on Open Collaboration*, p. 7. ACM, 2015.
- [11] A. Kobren, C. H. Tan, P. Ipeirotis, and E. Gabrilovich, "Getting More for Less: Optimized Crowdsourcing with Dynamic Tasks and Goals," In *Proceedings of the 24th international conference on world wide web*, pp. 592–602, International World Wide Web Conferences Steering Committee, 2015.
- [12] T. de Vreede, C. Nguyen, G.-J. de Vreede, I. Boughzala, O. Oh, and R. Reiter-Palmon, "A Theoretical Model of User Engagement in Crowdsourcing," in *Collaboration and Technology*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 94–109.
- [13] C. D. Nguyen, N. Tahmasbi, T. De Vreede, G.-J. De Vreede, O. Oh, and R. Reiter-Palmon, "Participant Engagement in Community Crowdsourcing," in *ECIS*, 2015.
- [14] U. av, "Uppdrag att redovisa fördjupade analyser om regelefterlevnaden inom yrkestrafiken på väg," Swedish Transport Agency, 2014.
- [15] H. Sternberg and B. Lantz, "Using crowdsourced data to analyze transport crime," *International Journal of Logistics Research and Applications*, vol. 21, no. 2, pp. 133–147, Jan. 2018.
- [16] R. J. Vallerand, L. G. Pelletier, M. R. Blais, N. M. Briere, C. Senecal, and E. F. Vallieres, "The Academic Motivation Scale: A Measure of Intrinsic, Extrinsic, and Amotivation in Education," *Educational and Psychological Measurement*, vol. 52, no. 4, pp. 1003–1017, Dec. 1992.
- [17] N. Kaufmann, T. Schulze, and D. Veit, "More than fun and money. Worker Motivation in Crowdsourcing-A Study on Mechanical Turk.," in *AMCIS*, 2011, vol. 11, pp. 1–11.
- [18] H. Liang, M.-M. Wang, J.-J. Wang, and Y. Xue, "How intrinsic motivation and extrinsic incentives affect task effort in crowdsourcing contests: A mediated moderation model," *Computers in Human Behavior*, vol. 81, pp. 168–176, Apr. 2018.
- [19] J. R. Hackman and G. R. Oldham, "Motivation through the design of work: Test of a theory," *Organizational Behavior and Human Performance*, vol. 16, no. 2, pp. 250–279, 1976.
- [20] N. Kaufmann and T. Schulze, "Worker Motivation in Crowdsourcing and Human Computation," *Education*, vol. 17, p. 6, 2011.
- [21] F. D. Davis, "A technology acceptance model for empirically testing new end-user information systems: Theory and results," PhD Thesis, Massachusetts Institute of Technology, 1985.
- [22] J. Pedersen *et al.*, "Conceptual Foundations of Crowdsourcing: A Review of IS Research." In *System Sciences (HICSS)*, 2013 46th Hawaii International Conference, pp. 579–588, IEEE, 2013.
- [23] S. Deterding, D. Dixon, R. Khaled, and L. Nacke, "From game design elements to gamefulness: defining 'gamification,'" presented at the MindTrek'11, Tampere, Finland, 2011.
- [24] J. Looyestyn, J. Kernot, K. Boshoff, J. Ryan, S. Edney, and C. Maher, "Does gamification increase engagement with online programs? A systematic review," *PLOS ONE*, vol. 12, no. 3, p. e0173403, Mar. 2017.
- [25] M. Sigala, "Gamification for Crowdsourcing Marketing Practices: Applications and Benefits

- in Tourism,” *Advances in Crowdsourcing*, pp. 129–145, 2015.
- [26] A. Dumitrache, L. Aroyo, C. Welty, R.-J. Sips, and A. Levas, “‘Dr. Detective’: combining gamification techniques and crowdsourcing to create a gold standard in medical text,” 2013.
- [27] K. Dergousoff and R. L. Mandryk, “Mobile Gamification for Crowdsourcing Data Collection: Leveraging the Freemium Model,” presented at the Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, 2015, pp. 1065–1074.
- [28] J. Hamari, J. Koivisto, and H. Sarsa, “Does Gamification Work? -- A Literature Review of Empirical Studies on Gamification.” In *System Sciences (HICSS), 2014 47th Hawaii International Conference*, pp. 3025–3034, IEEE, 2014.
- [29] B. Morschheuser, J. Hamari, J. Koivisto, and A. Maedche, “Gamified crowdsourcing: Conceptualization, literature review, and future agenda,” *International Journal of Human-Computer Studies*, vol. 106, pp. 26–43, Oct. 2017.
- [30] F. D. Davis, “Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology,” *MIS Quarterly*, vol. 13, no. 3, p. 319, Sep. 1989.
- [31] M. Lips-Wiersma and S. Wright, “Measuring the Meaning of Meaningful Work: Development and Validation of the Comprehensive Meaningful Work Scale (CMWS),” *Group & Organization Management*, vol. 37, no. 5, pp. 655–685, Oct. 2012.
- [32] M. F. Steger, B. J. Dik, and Y. Shim, “Assessing meaning and satisfaction at work,” *The Oxford Handbook of Positive Psychology Assessment*, vol. 2, 2010.
- [33] Y. M. Guo and M. S. Poole, “Antecedents of flow in online shopping: a test of alternative models,” *Information Systems Journal*, vol. 19, no. 4, pp. 369–390, Jul. 2009.
- [34] C. Yoon and E. Rolland, “Knowledge-sharing in virtual communities: familiarity, anonymity and self-determination theory,” *Behaviour & Information Technology*, vol. 31, no. 11, pp. 1133–1143, Nov. 2012.
- [35] R. F. Adler, D. Rubin, A. R. Mohammad, A. Irfan, H. Senadeera, and T. Nguyen, “That’s Easy! The Effects of Objective and Subjective Task Difficulty when Multitasking,” *Procedia Manufacturing*, vol. 3, pp. 5467–5474, 2015.
- [36] R. F. DeVellis, *Scale development: Theory and applications*, vol. 26. Sage publications, 2016.
- [37] A. B. Costello and J. W. Osborne, “Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most from Your Analysis,” *Exploratory Factor Analysis*, vol. 10, no. 7, p. 9, 2005.
- [38] J. S. Hickman and E. S. Geller, “Self-Management to Increase Safe Driving Among Short-Haul Truck Drivers,” *Journal of Organizational Behavior Management*, vol. 23, no. 3, pp. 1–20, 2005.
- [39] J. P. Saldanha, C. Shane Hunt, and J. E. Mello, “Driver management that drives carrier performance,” *Journal of Business Logistics*, vol. 34, no. 1, pp. 15–32, 2013.

Appendix: Constructs, Items, and Scales

Construct (Factor)	Questionnaire Item	Item Indicator	Scale	Source
Personal interest and attitude (PIA)	I have a good sense of what makes these tasks meaningful.	Meaningful	5- point Likert scale of absolutely untrue to absolutely true	[32]
	I know this work makes positive difference in people's lives.	PosDiff	5- point Likert scale of absolutely untrue to absolutely true	[32]
	I contributed to services that enhance human well-being and/or the environment.	WellBeing	5- point Likert scale of absolutely untrue to absolutely true	[33]
	I spent a lot of time on things that are truly important.	Important	5- point Likert scale of absolutely untrue to absolutely true	[33]
Received feedback and perceived autonomy (RFA)	It was really clear to me that I was doing well.	DoingWell	7-point Likert scale of strongly disagree to strongly agree	[34]

	I was aware of how good/bad I was performing.	Perform	7-point Likert scale of strongly disagree to strongly agree	[34]
	I was free to express my ideas and opinions in this virtual community.	ExpressIdea	7-point Likert scale of strongly disagree to strongly agree	[34], [35]
	I felt in total control of my action.	Control	7-point Likert scale of strongly disagree to strongly agree	[34], [35]
Perceived goal clarity and ease of use (PEU)	The goals of the study were clearly defined.	GoalDefined	7-point Likert scale of strongly disagree to strongly agree	[34]
	My interaction with Cabotagestudien was clear and understandable.	ClearInteract	7-point Likert scale of likely to unlikely	[30]
	It was easy for me to become skillful at using Cabotagestudien.	Skillful	7-point Likert scale of likely to unlikely	[30]
	I found Cabotagestudien easy to use.	EaseUse	7-point Likert scale of likely to unlikely	[30]
Perceived required effort for the task (PET)	I found this to be a complex task.	Complex	7-point Likert scale of strongly disagree to strongly agree	[36]
	This task was mentally demanding.	Demanding	7-point Likert scale of strongly disagree to strongly agree	[36]
	This task required a lot of thought and problem solving.	Thoughts	7-point Likert scale of strongly disagree to strongly agree	[36]
	I found this to be a challenging task.	Challenging	7-point Likert scale of strongly disagree to strongly agree	[36]
	I found this to be a time-consuming task.	TimeConsum	7-point Likert scale of strongly disagree to strongly agree	[36]
Experienced joy from gamification (EJG) (reversed scale)	Receiving a new badge (on the leaderboard) was fun.	NewBadge	5- point Likert scale of absolutely untrue to absolutely true	Gaming elements adapted from [29]
	It was always exciting to see my name in the leaderboard.	TopLeader	5- point Likert scale of absolutely untrue to absolutely true	Gaming elements adapted from [29]
	I enjoyed earning points (points for tagging cars.	Points	5- point Likert scale of absolutely untrue to absolutely true	Gaming elements adapted from [29]
	I liked the chance of receiving sponsorship reward.	Reward	5- point Likert scale of absolutely untrue to absolutely true	Gaming elements adapted from [29]